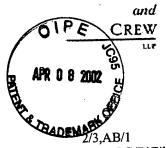
TOWNSEND and TOWNSEND

## DIALOG® WORLD PATENT INDEX SEARCH FOR ENGLISH LANGUAGE ABSTRACT OF EP 0875567 A2 COMPLETED 03/27/02 FOR 19496-005830 BY BRAD J. LOOS



DIALOG(R)File 351:Derwent WPI (c) 2002 Derwent Info Ltd. All rts. reserv.

012142521

WPI Acc No: 1998-559433/ 199848 XRAM Acc No: C98-167613 XRPX Acc No: N98-436229

Myc-binding zinc finger protein - useful for identifying transcription

modulating substances

Patent Assignee: BASF AG (BADI ); PROLIFIX LTD (PROL-N)

Inventor: EILERS M; HAENEL F; PEUKERT K Number of Countries: 027 Number of Patents: 004

Patent Family:

Patent No Kind Date Applicat No Kind Date Week
EP 875567 A2 19981104 EP 98106426 A 19980408 199848 B
DE 19718249 A1 19981105 DE 1018249 A 19970430 199850
JP 11001498 A 19990106 JP 98118863 A 19980428 199911
-US-6160091 A 20001212 US 9863035 A 19980421 200067

Priority Applications (No Type Date): DE 1018249 A 19970430

Patent Details:

Patent No Kind Lan Pg Main IPC Filing Notes

EP 875567 A2 G 13 C12N-015/12

Designated States (Regional): AL AT BE CH CY DE DK ES FI FR GB GR IE IT

LI LT LU LV MC MK NL PT RO SE SI

DE 19718249 A1 C07K-014/435

JP 11001498 A 21 C07K-014/47

US 6160091 A C07K-014/47

Abstract (Basic): EP 875567 A

A protein (SEQ 2) (a defined sequence of 803 amino acids given in the specification) is new.

USE - The organisms can be used to produce the protein (a Myc-binding zinc finger protein). The protein can be used to identify transcription-modulating substances, in a process comprising: (a) incubating the protein with myc gene product under conditions such that a complex is formed between the two proteins; (b) repeating (a) in the presence of one or more test substances; (c) determining the difference in protein complex formation between (a) and (b); and (d) selecting substances for which the protein complex formation in (b) is different from that of (a). The protein can be used to produce antibodies. The nucleic acid sequence can be used for gene therapy. The nucleic acid sequence complementary to the above sequence is useful for gene therapy (claimed).

Dwg.0/0

PA 3210859 v1



**Europäisches Patentamt** 

**European Patent Office** 

Office européen des brevets



11) EP 0 875 567 A2

(12)

### **EUROPÄISCHE PATENTANMELDUNG**

- (43) Veröffentlichungstag: 04.11.1998 Patentblatt 1998/45
- (21) Anmeldenummer: 98106426.4
- (22) Anmeldetag: 08.04.1998

- (51) Int. Cl.<sup>6</sup>: **C12N 15/12**, C07K 14/47, C12N 15/63, C12N 1/21, G01N 33/68, C07K 16/18, A61K 48/00
- (84) Benannte Vertragsstaaten:
  AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU
  MC NL PT SE
  Benannte Erstreckungsstaaten:
  AL LT LV MK RO SI
- (30) Priorität: 30.04.1997 DE 19718249
- (71) Anmelder:
  BASF AKTIENGESELLSCHAFT
  67056 Ludwigshafen (DE)

- (72) Erfinder:
  - Peukert, Karen 35094 Lahntal-Sterzhausen (DE)
  - Haenel, Frank, Dr. 07745 Jena (DE)
  - Eilers, Martin, Prof. Dr. 35043 Marburg-Cappel (DE)
- (54) Myc-bindende Zinkfinger-Proteine, ihre Herstellung und ihre Verwendung
- (57) Neue Myc-bindende Zinkfingerproteine, ihre Herstellung und ihre Verwendung.

### Beschreibung

10

Die vorliegend Erfindung betrifft Myc-bindende Zinkfinger-Proteine, ihre Herstellung und ihre V rwendung.

Myc ist ein spezifisch an DNA bindendes Protein. Es wird zur Familie der Helix-Loop-Helix/Leucin-Zipper (HLH/LZ) Transkriptionsfaktoren gezählt (Landschulz et al., 1988, Murre et al., 1989). Myc ist ein zentraler Transkriptionsaktivator, der mit dem Protein Max (Amati et al., 1993) einen Komplex bildet und durch diesen molekularen Mechanismus andere Gene aktiviert, beispielsweise alpha-Prothymosingen, Ornithindecarboxylasegen und cdc25A.

Von Schulz et al, 1995, wurde ein 13 Zinkfinger enthaltendes Protein aus der Maus beschrieben, dessen zelluläre

Funktion jedoch unklar ist.

Aufgrund seiner Schlüsselstellung in der Transkription bietet Myc einen Ansatzpunkt zum Verständnis von zellulären, insbesondere von pathophysiologischen Prozessen.

Es bestand daher die Aufgabe, weitere Informationen über die molekulare Wirkungsweise von Myc, insbesondere über die Myc vermittelte Genrepression bereitzustellen.

Gegenstand der Erfindung ist ein Protein mit der in SEQ ID NO:2 dargestellten Aminosäuresequenz. Dieses Pro-

tein besitzt dreizehn Zinklingerdomänen.

Es weist folgende biologischen Eigenschaften auf:

Spezifische Bindung an Myc,

- · Transaktivierung des Adenovirus Major Late (AdML) Promotors,
- 20 Transaktivierung des Cyclin D1 Promotors,
  - durch Assoziation mit Myc wird die Transaktivierung gehemmt,

in Abwesenheit von Myc ist das Protein im wesentlichen im Cytosol assoziiert mit Mikrotubuli zu finden.

Ein weiterer Gegenstand der Erfindung sind Proteine, die sich aus der SEQ ID NO:2 dargestellten Struktur Erch Substitution, Insertion oder Deletion von einem oder mehreren Aminosauren ableiten lassen, wobei diese Proteine noch die wesentlichen biologischen Eigenschaften des durch SEQ ID NO:2 beschriebenen Proteins besitzen. Dese Proteine werden im folgenden Muteine genannt. Unter wesentlichen Eigenschaften wird die spezifische Bindung der Muteine an Myc verstanden.

Die oben aufgeführten Eigenschaften des durch SEQ ID NO:2 beschriebenen-Proteins-müssen nicht alle bei den Muteinen vorhanden sein, solange die spezifische Bindung an Myc gegeben ist. Bevorzugt sind jedoch diejenigen Muteine, die alle der oben aufgeführten Eigenschaften besitzen.

Die Anzahl der durch Insertion Substitution oder Deletion gegenüber dem durch SEQ ID NO:2 beschriebenen Protein veränderten Aminosäuren kann zwischen 1 und 100, bevorzugt zwischen 1 und 50 Aminosäuren variieren. Die Veränderungen können in einem kleineren Bereich des Moleküls konzentriert oder auch über das ganze Molekül verteilt sein.

Bevorzugte Veränderungen sind konservative Substitutionen, bei denen eine Aminosäure durch eine andere Aminosäure mit ähnlicher Raumerfüllung, Ladung oder Hydrophilie ersetzt wird.

Beispiele für solche konservativen Substitutionen sind

Ersatz von Arg durch Lys oder umgekehrt,
Ersatz von Arg durch His oder umgekehrt,
Ersatz von Asp durch Glu oder umgekehrt,
Ersatz von Asn durch Gln oder umgekehrt,
Ersatz von Cys durch Met oder umgekehrt,
Ersatz von Cys durch Ser oder umgekehrt,

Ersatz von Gly durch Ala oder umgekehrt, Ersatz von Val durch Leu oder umgekehrt, Ersatz von Val durch IIe oder umgekehrt, Ersatz von Leu durch IIe oder umgekehrt,

Ersatz von Phe durch Tyr oder umgekehrt, Ersatz von Phe durch Trp oder umgekehrt, Ersatz von Ser durch Thr oder umgekehrt.

Die Veränderungen können auch kombiniert werden, z.B. eine oder mehrere Substitutionen mit Deletionen und/oder Insertionen.

Ein weiterer Gegenstand der Erfindung sind Nukleinsäuresequenzen, die für die oben beschriebenen Proteine codieren. Solche Nukleinsäuresequenzen sind bevorzugt DNA, insbesondere cDNA Sequenzen, in einz Isträngig r oder doppelsträngiger Form.

Bevorzugte Nukl insäuresequenzen sind solche mit der in SEQ ID NO:1 dargestellt in Sequenz und solche, die mit dieser Sequenz einen hohen Verwandschaftsgrad aufweisen, beispielsw ise solche, die für das gleiche Protein codieren wie SEQ ID NO:1. Weitere bevorzugte Nukleinsäuresequenzen sind solche, die für ein Protein codieren, das 95% oder mehr Identität mit dem Protein der Sequenz SEQ ID NO:2 aufweist.

Ein weiterer Gegenstand der Erfindung sind Vektoren, die eine der oben beschriebenen Nukleinsäuresequenzen in funktioneller Verknüpfung mit einem oder mehreren Regulationselementen tragen. Unter Regulationselemente sind Nukleinsäurefragmente zu verstehen, die auf Transkription oder Translation einen regulierenden Einfluß haben, beispielsweise Promotoren, Enhancer, Polyadenylierungsstellen, ribosomale Bindungsstellen.

Die mit solchen Vektoren transformierten Wirtsorganismen sind ebenfalls Gegenstand der Erfindung. Als Wirtsorganismen geeignet sind Mikroorganismen, pflanzliche oder tierische Zellen oder Lebewesen. Bevorzugte Wirtsorganismen sind eukaryontische Zellen und Lebewesen. Der Begriff Wirtsorganismus umfaßt auch beispielsweise transgene Tiere und Pflanzen.

Die Herstellung der erfindungsgemäßen Proteine erfolgt bevorzugt mit Hilfe gentechnischer Verfahren. Ein Wirtsorganismus, der die Erbinformation für die erfindungsgemäßen Proteine trägt, wird unter Bedingungen kultiviert, die die
Expression des Proteins erlauben. Diese Bedingungen -wie Temperatur, Nährmedium, Zelldichte - hängen weitgehend
von der Wahl des Wirtsorganismus ab. Solche Bedingungen sind jedoch dem Fachmann für die einzelnen Wirtsorganismen geläufig.

Die exprimierten Proteine werden anschließend, ggf. nach Aufbrechen des Wirtsorganismus, vom Wirtsorganismus abgetrennt und in reiner Form durch bekannte Methoden der Proteinreinigung, wie Fällung, Chromatographie, Elektrophorese in reiner Form isoliert. Ein weiterer Gegenstand der Erfindung ist die Verwendung der Proteine als Antigen zur Herstellung von Antikörpern, sowie die so erhaltenen Antikörper. Es lassen sich durch dem Fachmann bekannte Verfahren polyklonale Antiseren oder auch monoklonale Antikörper herstellen.

Die erfindungsgemäßen Proteine eignen sich auch als Testsysteme zur Auffindung von potentiellen selektiven Transkriptionsmodulierenden Substanzen. Dies läßt sich besonders gut testen, indem man die Fähigkeit der Proteine, mit Myc einen Proteinkomplex zu bilden, ausnützt. Ein weiterer Gegenstand der Erfindung ist daher ein Verfahren zur Identifizierung von spezifischen transkriptionsmodulierenden Substanzen, das folgende Schritte umfaßt:

- (a) Inkubation des Proteins gemäß Anspruch 1 mit dem Genprodukt von myc unter Bedingungen, unter denen sich ein Proteinkomplex zwischen diesen beiden Proteinen ausbildet,
- (b) Inkubation der beiden Proteine unter ansonst gleichen Bedingungen wie (a) jedoch in Anwesenheit einer oder mehrerer Substanzen, die auf spezifische transkriptionsmodulierende Aktivitäten zu testen sind,
- (c) Ermitteln des Unterschiedes in der Proteinkomplexbildung zwischen (b) und (a),
- (d) Auswahl solcher Substanzen, bei denen gemäß Schritt (b) eine andere Proteinkomplexbildung erhalten wurde als bei Schritt (a).

Es lassen sich damit Substanzen auffinden, die die Proteinkomplexbildung zwischen den neuen Zinkfingerprotein und Myc fördern, aber auch solche, die sie unterbinden.

Die erfindungsgemäßen Nukleinsäuresequenzen eignen sich auch zur Gentherapie von Erkrankungen, bei denen die durch Myc vermittelte Transkription gestört ist.

Beispielsweise können zusätzliche Gensequenzen eingebracht werden um so die zelluläre Konzentration der Zinkfingerproteine zu erhöhen. Es kann aber auch gewünscht sein, daß die Konzentration der Zinkfingerproteine erniedrigt werden soll. In diesem Falle bietet sich eine Gentherapie auf antisense Basis an, wobei man eine zu dem Zinkfingerproteingen komplementäre Nukleinsäure oder Nukleinsäurederivat appliziert, und somit die Expression des Zinkfingerproteingens reduziert.

Die weitere Ausgestaltung der Erfindung ist in den folgenden Beispielen aufgeführt.

### 50 Beispiel 1

30

35

Isolierung der DNA mit der durch SEQ ID NO:1 beschriebenen Struktur

Vorausgegangene Arbeiten hatten gezeigt, daß die Integrität der Helix-Loop-Helix Domäne von Myc kritisch für die Genrepression durch Myc in stabilen Zellinien war (Philipp et al., 1994). Um neue Proteine zu identifizieren, die mit dem C-Terminus von Myc interagieren, wurde ein DNA-Fragment, das für die basische Region und die HLH/LZ Domäne (Aminosäuren 355-439 des humanen Myc) codiert, im Leserahmen an die DNA bindend Domäne von GAL4 (Aminosäuren 1-147) fusioniert und als Köder in inem "Two-Hybrid-Screen" (Fields and Song, 1989) benutzt.

2x10<sup>5</sup> unabhängige Transformanden einer HeLa cDNA Bibliothek, markiert mit der GAL4 Aktivierungsdomäne, wurden gescreent. Ein Clon mit β-Galaktosidaseaktivität wurde weiter charakterisiert. Es wurde keine Interaktion zwischen dem von diesem Clon codierten Protein und der DNA Bindungsdomäne von GAL4 allein oder einer GAL4-BCY-1 Chimäre, die als Negativkontrolle benutzt wurde, festgestellt.

Die Interaktion mit Myc wurde aufgehoben durch Deletion der HLH-Domäne in Myc (370-412), nicht aber durch Insertion der vier Aminosäuren zwischen der HLH Domäne und dem Leucin-Zipper (In 412) oder durch Deletion des gesamten Leucin-Zippers (412-434). Eine spezifische Interaktion wurde auch nachgewiesen mit N-Myc aber keine mit MAX oder USF, zwei HLH-Proteinen, die mit Myc nahe verwandt sind.

cDNA-Moleküle mit voller Länge wurden durch ein 5'-RACE-Protokoll isoliert und sequenziert (SEQ ID NO:1). Sie codieren ein Protein mit 803 Aminosäuren (SEQ ID NO:2) mit einem theoretischen Molekulargewicht von 87,970 Dalton. Das Protein wurde Miz-1 für Myc-Interacting-Zincfinger-Protein-1 genannt.

Die Sequenzierung ergab, daß der isolierte Clon für ein Zinkfingerprotein mit 13 Zinkfingern codierte, 12 davon unmittelbar geclustert in der C-terminalen Hälfte des Proteins.

### 15 Beispiel 2

#### Herstellung von Muteinen

Ausgehend von der in SEQ ID NO:1 dargestellten Nukleinsäuresequenz können mit dem Fachmann geläufigen Methoden der Gentechnik Nukleinsäuren hergestellt werden, die für veränderte Proteine (Muteine) codieren. Die Herstellung der Muteine selbst erfolgt zweckmäßigerweise durch Expression einer Nukleinsäure in einem geeigneten Wirtsorganismus.

### Beispiel 3

25

50

Assoziation des Proteins SEQ ID NO:2 mit Myc

Der G-Terminus des Proteins SEQ ID NO:2 (Aminosaure 269-803) wurde mit der Glutathion-Transferase (GST) (Smith and Johnson, 1988) fusioniert, das GST-Miz-1 Fusionsprotein gereinigt und mit in vitro synthetisiertem, radioaktiv markiertem Myc Protein inkubiert. Myc assoziiert spezifisch mit GST-Miz-1, jedoch nicht mit GST. Eine Mutante von Myc, der die HLH Domäne fehlt, konnte nicht mit GST-Miz-1 assoziieren. Radioaktiv markiertes Max interagiert weder mit GST-Miz-1 noch mit GST. Jedoch kann mit Hilfe von Myc Max an GST-Miz-1-Kügelchen in vitro binden, was dafür spricht, daß Miz-1 und Max mit unterschiedlichen Flächen der HLH-Domäne von Myc interagieren.

### 35 Literaturverzeichnis

Amati, B., Brooks, M. W., Levy, N., Littlewood, T. D., Evan, G. I., and Land, H. (1993). Oncogenic activity of the c-Myc protein requires dimerization with Max. Cell 72, 233-245.

Fields, S., and Song, O. (1989). A novel genetic system to detect protein-protein interactions. Nature 340, 245-246.

Landschulz, W. H., Johnson, P. F., and McKnight, S. L. (1988). The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. Science 240, 1759-1764.

Murre, C., SchonleberMcCaw, P., and Baltimore, D. (1989). A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. Cell 56, 777-783.

Philipp, A., Schneider, A., Väsrik, I., Finke, K., Xiong, Y., Beach, D., Alitalo, K., and Eilers, M. (1994). Repression of Cyclin D1: a Novel Function of MYC. Mol. Cell. Biol. 14, 4032-4043.

Schulz, T. C., Hopwood, B., Rathjen, P. D., and Wells, J. R. (1995). An unusual arrangement of 13 zinc fingers in the vertebrate gene Z13. Biochem. J. 311, 219-224.

Smith, D. B., and Johnson, K. S. (1988). Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione-S-transferase. Gen 67, 31-40.

## SEQUENZ PROTOKOLL

	(1) ALGE	MEINE INFORMATION:	
10	(i)	ANMELDER:  (A) NAME: BASF Aktiengesellschaft  (B) STRASSE: Carl-Bosch-Strasse 38  (C) ORT: Ludwigshafen  (E) LAND: Bundesrepublik Deutschland  (F) POSTLEITZAHL: D-67056  (G) TELEPHON: 0621/6048526  (H) TELEFAX: 0621/6043123  (I) TELEX: 1762175170	
15	(ii)	ANMELDETITEL: Myc-bindende Zinkfingerproteine	
15	(iii)	ANZAHL DER SEQUENZEN: 2	
20	(iv)	COMPUTER-LESBARE FORM:  (A) DATENTRĀGER: Floppy disk  (B) COMPUTER: IBM PC compatible  (C) BETRIEBSSYSTEM: PC-DOS/MS-DOS  (D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPA)	
	(2) INFO	RMATION ZU SEQ ID NO: 1:	
25	(i)	SEQUENZ CHARAKTERISTIKA:  (A) LÄNGE: 2680 Basenpaare  (B) ART: Nukleinsäure  (C) STRANGFORM: Einzel	
30		(D) TOPOLOGIE: linear	
	(ii)	ART DES MOLEKÜLS: CDNS zu mRNS	
	(iii)	HYPOTHETISCH: NEIN	
35	(iii)	ANTISENSE: NEIN	
	(ix)	MERKMALE: (A) NAME/SCHLÜSSEL: 5'UTR (B) LAGE: 1159	
40	(ix)	MERKMALE: (A) NAME/SCHLÜSSEL: CDS (B) LAGE: 1602571	
45	(ix)	MERKMALE: (A) NAME/SCHLÜSSEL: 3'UTR (B) LAGE: 25722680	
	(xi)	SEQUENZBESCHREIBUNG: SEQ ID NO: 1:	
50	GGAGTGCC	COT CCCCGGCCTT CTCGCGGCCG TGATGCACCT CCCTCTGCGG TGGGGTCCGG	60
	<b>ሪአርአጥርርር</b>	AC CTAATGAGCC GGACGAGGGG AGCCAAGCTG GAGTTTACAC AGGCAAACTG	120

	TCAC	SAAA	AGA (	GTAG(	CTG	GG C	rgtc'	rgga	A AT	CTGA				rrr ( Phe F				174
5			~~~	C) III	cmc	mm/c	CAA	CAC	כתכ	አልሮ	CAG	CAG	ccc	CAG	· ርጥር	GGG		222
																Gly		
	HIS	ser	GIII	1112	10	Dea	014	<b></b>		15			5		20	•		
																		050
10														TTT				270
	Leu	Leu	Cys		Cys	Thr	Phe	Val		Asp	GIY	vaı	HIS	Phe 35	гÀ2	Ald		
				25					30					33				
	CAT	AAA	GCA	GTG	CTG	GCG	GCC	TGC	AGC	GAG	TAC	TTC	AAG	ATG	CTC	TTC		318
15	His	Lys	Ala	Va1	Leu	Ala	Ala	Cys	Ser	Glu	Tyr	Phe	Lys	Met	Leu	Phe		
	•		40					45					50					
	CTC	GAC	CAG	AAG	GAC	GTG	GTG	CAC	CTG	GAC	ATC	AGT	AAC	GCG	GCA	GGC		366
	Val	ASD	Gln	Lvs	Asp	Val	Val	His	Leu	Asp	Ile	Ser	Asn	Ala	Ala	Gly		
20		55		-			60					65						
20	*					a. a		1 mg	ma C	100	ccc	220	CTTC	» CC	CTIC	AGC		414
	CTG	GGG	CAG	ATG	CTG	GAG	TTT	Mot	TAC	Thr	Δla	LVS	Len	AGC Ser	Leu	Ser		
	ьеи 70	GIY	GIn	Met	pen	75	FIIE	Mec	TYT	1111	80	2,0	200			85		
25	CCT	GAG	AAC	GTG	GAT	GAT	GTG	CTG	GCC	GTG	GCC	ACT	TTC	CTC	CAA	ATG		462
	Pro	Glu	Asn	Val		Asp	Val	Leu	Ala		Ala	Thr	Phe	Leu		Met		
					90					95					100			
	CAG	GAC	ATC	ATC	ACG	GCC	TGC	CAT	GCC	CTC	AAG	TCA	CTT	GCT	GAG	-ccg-		-510
30														Ala				
				105					110					115				
	CCM	» CC	NGC.	CCT	GGG	GGA	аат	GCG	GAG	GCC	TTG	GCC	ACA	GAA	GGA	GGG	•	558
	Ala	whr.	Ser	Pro	Glv	Glv	Asn	Ala	Glu	Ala	Leu	Ala	Thr	Glu	Gly	Gly		
35	AI G		120					125					130					
									ama	ccc	3.00	»cc	N ČG	CTC	»GC	AGG		606
	GAC	AAG -	AGA	GCC	AAA	GAG	GAG	AAG	GIG Wal	Ala	Thr	Ser	Thr	CTG Leu	Ser	Ara		000
	_		Arg	Ala	рух		140	пÃЭ	Vai	AIG	1111	145				;		
40		135													1.		,	
	CTG	GAG	CAG	GCA	GGA	CGC	AGC	ACA	CCC	ATA	GGC	CCC	AGC	AGG	GAC	CTC		654
			Gln										ser	Arg	Asp			
	150	•				155					100			<i>:</i>		165		
45	AAG	GAG	GAG	CGC	GGC	GGT	CAG	GCC	CAG	AGT	GCG	GCC	AGC	GGT	GCA	GAG		702
,,,	Lys	Glu	Glu	Arg	Gly	Gly	Gln	Ala	Gln	Ser	Ala	Ala	Ser	Gly	Ala	Glu		
	-			٠	170					175				•	180		•	
	C3.C	3.03	CAC	272	פכר	Сът	GCG	רכר	CGG	GAG	CCG	CCG	CCT	GTG	GÄG	CTC		750
														Val				
50	GIII	1 111	,	185		2			190					195				

										oom.	~~~	001	C	ccm	CAC	ccc	CCT	798
		AAG	CCA	GAC	CCC	ACG	AGT	GGC	AIG	GCT	330	BLA	Clu	772	Clu	Ala	Δla	,,,,
		Lys	Pro		Pro	Thr	ser			Ald	Ald	Ald	GIU	210	GIU	VIG	ALG	
	5			200					205					210				
		TTG	ጥሮሮ	CAC	AGC	TCG	GAG	CAA	GAA	ATG	GAG	GTG	GAG	CCC	GCC	CGG	AAA	846
		Leu	Sor	Clu	Ser	Ser	Glu	Gln	Glu	Met	Glu	Va1	Glu	Pro	Ala	Arg	Lys	
			215	GIU.	001	<b>D</b> 01		220			_		225					
	10	GGG	GAA	GAG	GAG	CAA	AAG	GAG	CAA	GAG	GAG	CAA	GAG	GAG	GAG	GGC	GCA	894
		Gly	Glu	Glu	Glu	Gln	Lys	Glu	Gln	Glu	Glu	Gln	Glu	Glu	Glu	Gly	Ala	
		230		,			235					240					245	•
												~~~	áma		330	CCN	CAC	942
	•	GGG	CCA	GCT	GAG	GTC	AAG	GAG	GAG	GGT	TCC	CAG	CTG	GAG	AAC	Clv	Clu	, ,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,
	15	Gly	Pro	Ala	Glu	Val	Lys	Glu	Giu	GIA		GIN	ren	GIU	ASII	260	ĢIU	
•						250					255					200		
•		000	ccc	CAC	CAG	AAC	CAC	ልልጥ	GAG	GAG	TCA	GCG	GGC	ACA	GAC	TCG	GGG	990
		33-	D-0	CAU	Clu	Asn	Glu	Asn	Glu	Glu	Ser	Ala	Glv	Thr	Asp	Ser	Gly	
	20	Ата	PIO		265	ASII	GIU	non	014	270					275			
	20																	
		CAG	GAG	CTC	GGC	TCC	GAG	GCC	CGG	GGC	CTG	CGC	TCA	GGC	ACC	TAC	GGC	1038
		Gln	Glu	Leu	Gly	Ser	Glu	Ala	Arg	Gly	Leu	Arg	Ser	Gly	Thr	Tyr	GLY	
				280					285					290				
	25								_					~~~		mee	CAC	1086
		GAC	CGC	ACG	GAG	TCC	AAG	GCC	TAC	GGC	TCC	GTC	ATC	CAC	AAG	TGC Cvc	Clu	1000
		Asp	Arg	Thr	Glu	Ser	ГЛS		Tyr	Gly	ser	vaı		HIS	гÃЯ	Cys	GIU	
			295					300					305				•	
		CAC	ጥርጥ	ccc	AAG	GAG	ጥጥር	ACG	CAC	ACG	GGG	AAC	TTC	AAG	CGG	CAC	ATC	1134
	30	JAC	Care	Clv	Lyg	Glu	Phe	Thr	His	Thr	Glv	Asn	Phe	Lys	Arg	His	Ile	
		310	Cys	GIY	פעם	014	315					320					325	
												-						
		CGC	ATC	CAC	ACG	GGG	GAG	AAG	CCC	TTC	TCG	TGC	CGG	GAG	TGC	AGC	AAG	1182
	35	Arg	Ile	His	Thr	Gly	Glu	Lys	Pro	Phe	Ser	Cys	Arg	Glu	Сув	Ser	Lys	
						330					335					340		
								000	mac	330	CCC	CAM	CAC	AAG	ACG	CAC	AGC	1230
		GCC	TTT	TCC	GAC	CCG	GCC	GCG	TGC	AAG	. GCC	LAI	GAG	Tare	Thr	His	AGC	
		Ala	Phe	Ser			Ala	Ala	Cys	350		1113		2,5	355		Ser	
	40				345	1												
		ርርጥ	CTG	AAG	.ccc	TAC	GGC	TGC	GAG	GAG	TGC	GGG	AAG	AGC	TAC	CGC	CTC	1278
		Pro	Leu	Lvs	Pro	Tyr	Gly	Cys	Glu	G1 u	Суя	Gly	Lys	Ser	Tyr	Arg	Leu	
				360			-		365					370				
4																		
Դ. •	45																	1276
•	.45	ATC	AGC	CTG	CTC	AAC	CTG	CAC	AAG	AAC	CGG	CAC	TCG	GGC	GAG	GCC	CGC	1326
75. - 	45	ATC	AGC Ser	CTG	CTC	AAC Asn	CTG Leu	His	Lys	AAG Lys	CGG Arg	CAC His	Ser	Gly	GAG Glu	GCG Ala	CGC	1326
•	.45	ATC Ile	AGC Ser	CTG	CTC Lev	AAC Asn	CTG Leu	CAC His	Lys	AAC Lys	CGG Arg	CAC His	TCG Ser 385	Gly	GAG Glu	GCC Ala	CGC Arg	1326
•	45	Ile	Ser 375	CTG Leu	Leu	Asn	Leu	His 380	Lys	Lys	Arg	, His	385	Gly	Glu	Ala	Arg	1326 1374
•	.45 50	Ile TAC	Ser 375	CTG Leu	Lev	Asn GAC	Leu TGC	His 380	Lys AAC	Lys CTC	Arg	His ACC	Ser 385	Gly TCC	Glu GGC	Ala	Arg	
•		Ile TAC	Ser 375 CGC	CTG Leu	Lev	Asn GAC	Leu TGC	His 380 GGC Gly	Lys AAC	Lys CTC	Arg	His ACC	385 ACC	Gly TCC	Glu GGC	Ala	Arg	

\_

•					CTG												1422
	Lys	Arg	His	GIn	Leu 410	vaı	HIS	ser	GIĀ	415	гÀ2	PIO	ığı	GIII	420	voħ	
5	m> C	maa	ccc	ccc	TCC	ጥጥር	ጥርር	GAC	כככ	аст	TCC	AAG	ATG	CGC	CAC	CTG	1470
					Ser												
	TYL	Cys	GIY	425	501		501		430					435			
10	GAG	ACC	CAC	GAC	ACG	GAC	AAG	GAG	CAC	AAG	TGC	CCA	CAC	TGC	GAC	AAG	1518
					Thr												
	•		440					445					450				
	AAG	TTC	AAC	CAG	GTA	GGG	AAC	CTG	AAG	GCC	CAC	CTG	AAG	ATC	CAC	ATC	1566
15	Lys	Phe	Asn	Gln	Val	Gly	ÀSN	Leu	Lys	Ala	His		Lys	Ile	His	Ile	
		455					460					465					
	GCT	GAC	GGG	CCC	CTC	AAG	TGC	CGA	GAG	TGT	GGG	AAG	CAG	TTC	ACC	ACC	1614
	Ala	Asp	Gly	Pro	Leu		Cys	Arg	Glu	Суѕ		Lys	Gln	Phe	Thr		
20	470					475					480					485	
	TCA	GGG	AAC	CTG	AAG	CGG	CAA	CTT	CGG	ATC	CAC	AGC	GGG	GAG	AAG	CCC	1662
	Ser	Gly	Asn	Leu	Lys	Arg	Gln	Leu	Arg	Ile	His	Ser	Gly	Glu	Lys	Pro	
					490					495					500		
25		ama	mcc	አመም	CAC	TICC	CAG	CGA	CAG	ጥጥጥ	GCA	GAC	CCC	GGC	GCT	CTG	1710
		616	166	MIC	CAC	100											
	m	17 - 1	Cve	Tle	His	Cvs	Gln	Arσ	Gln	Phe		Asp		Gly	Ala	Leu	
	Tyr	Val	Сув		His	Cys	Gln	Arg	Gln 510	Phe		Asp		Gly 515	Ala	Leu	
-	<u>-</u>			505					510		Ala		Pro	515			1758
 30	CAG	CGG	CAC	505 GTC	CGC	ATT	CAC	ACA	510 GGT	GAG	Ala AAG	CCA	Pro	515 CAG	TGT	GTG	1758
 30	CAG	CGG	CAC His	505 GTC		ATT	CAC	ACA Thr	510 GGT	GAG	Ala AAG	CCA	TGC Cys	515 CAG Gln	TGT	GTG	1758
30	CAG Gln	CGG Arg	CAC His 520	STC Val	CGC Arg	ATT Ile	CAC His	ACA Thr 525	GGT Gly	GAG Glu	Ala AAG Lys	CCA Pro	TGC Cys 530	CAG Gln	TGT Cys	GTG Val	
30	CAG Gln ATG	CGG Arg	CAC His 520	GTC Val	CGC Arg	ATT Ile	CAC His	ACA Thr 525 CAG	GGT Gly GCC	GAG Glu AGC	AAG Lys	CCA Pro	TGC Cys 530 ATC	CAG Gln GCC	TGT Cys CAC	GTG Val GTG	1758
30	CAG Gln ATG	CGG Arg	CAC His 520	GTC Val	CGC Arg	ATT Ile	CAC His	ACA Thr 525 CAG	GGT Gly GCC	GAG Glu AGC	AAG Lys	CCA Pro CTC Leu	TGC Cys 530 ATC	CAG Gln GCC	TGT Cys CAC	GTG Val GTG	
	CAG Gln ATG	CGG Arg	CAC His 520	GTC Val	CGC Arg	ATT Ile	CAC His	ACA Thr 525 CAG	GGT Gly GCC	GAG Glu AGC	AAG Lys	CCA Pro	TGC Cys 530 ATC	CAG Gln GCC	TGT Cys CAC	GTG Val GTG	
	CAG Gln ATG Met	CGG Arg TGC Cys 535	CAC His 520 GGT Gly	GTC Val	CGC Arg GCC Ala	ATT Ile TTC Phe	CAC His ACC Thr 540	ACA Thr 525 CAG Gln	GGT Gly GCC Ala	GAG Glu AGC Ser	Ala AAG Lys TCC Ser	CCA Pro CTC Leu 545	TGC Cys 530 ATC Ile	CAG Gln GCC Ala	TGT Cys CAC His	GTG Val GTG Val	
	CAG Gln ATG Met	CGG Arg TGC Cys 535	CAC His 520 GGT Gly	GTC Val AAG Lys	CGC Arg	ATT Ile TTC Phe	CAC His ACC Thr 540	ACA Thr 525 CAG Gln	GGT Gly GCC Ala	GAG Glu AGC Ser	Ala  AAG Lys  TCC Ser	CCA Pro CTC Leu 545	TGC Cys 530 ATC Ile	CAG Gln GCC Ala	TGT Cys CAC His	GTG Val GTG Val AAG Lys	1806
35	CAG Gln ATG Met	CGG Arg TGC Cys 535	CAC His 520 GGT Gly	GTC Val AAG Lys	CGC Arg GCC Ala	ATT Ile TTC Phe	CAC His ACC Thr 540	ACA Thr 525 CAG Gln	GGT Gly GCC Ala	GAG Glu AGC Ser	Ala  AAG Lys  TCC Ser	CCA Pro CTC Leu 545	TGC Cys 530 ATC Ile	CAG Gln GCC Ala	TGT Cys CAC His	GTG Val GTG Val	1806
	CAG Gln ATG Met CGC Arg 550	CGG Arg TGC Cys 535 CAG Gln	CAC His 520 GGT Gly CAC	GTC Val AAG Lys ACC	CGC Arg GCC Ala GGG Gly	ATT Ile TTC Phe GAG Glu 555	CAC His ACC Thr 540 AAG	ACA Thr 525 CAG Gln CCC Pro	GGT Gly GCC Ala TAC Tyr	GAG Glu AGC Ser GTC Val	Ala AAG Lys TCC Ser TGC Cys 560	CCA Pro CTC Leu 545 GAG Glu	TGC Cys 530 ATC Ile CGC Arg	CAG Gln GCC Ala TGC Cys	TGT Cys CAC His GGC Gly	GTG Val GTG Val AAG Lys 565	1806
35	CAG Gln ATG Met CGC Arg 550	CGG Arg TGC Cys 535 CAG Gln	CAC His 520 GGT Gly CAC His	GTC Val AAG Lys ACC Thr	CGC Arg GCC Ala GGG Gly	ATT Ile TTC Phe GAG Glu 555	CAC His ACC Thr 540 AAG Lys	ACA Thr 525 CAG Gln CCC Pro	GGT Gly GCC Ala TAC Tyr	GAG Glu AGC Ser GTC Val	Ala  AAG Lys  TCC Ser  TGC Cys 560 CAT	CCA Pro CTC Leu 545 GAG Glu	TGC Cys 530 ATC Ile CGC Arg	CAG Gln GCC Ala TGC Cys	TGT Cys CAC His GGC Gly	GTG Val GTG Val AAG Lys 565	1806 1854
35	CAG Gln ATG Met CGC Arg 550	CGG Arg TGC Cys 535 CAG Gln	CAC His 520 GGT Gly CAC His	GTC Val AAG Lys ACC Thr	CGC Arg GCC Ala GGG Gly	ATT Ile TTC Phe GAG Glu 555	CAC His ACC Thr 540 AAG Lys	ACA Thr 525 CAG Gln CCC Pro	GGT Gly GCC Ala TAC Tyr	GAG Glu AGC Ser GTC Val	Ala  AAG Lys  TCC Ser  TGC Cys 560 CAT	CCA Pro CTC Leu 545 GAG Glu	TGC Cys 530 ATC Ile CGC Arg	CAG Gln GCC Ala TGC Cys	TGT Cys CAC His GGC Gly	GTG Val GTG Val AAG Lys 565	1806 1854
35	CAG Gln ATG Met CGC Arg 550 AGA Arg	CGG Arg TGC Cys 535 CAG Gln TTC	CAC His 520 GGT Gly CAC His	GTC Val AAG Lys ACC Thr	CGC Arg GCC Ala GGG Gly TCC Ser 570	ATT Ile TTC Phe GAG Glu 555 AGC	CAC His ACC Thr 540 AAG Lys CAG	ACA Thr 525 CAG Gln CCC Pro	GCC Ala	GAG Glu AGC Ser GTC Val AAT Asn 575	Ala  AAG Lys  TCC Ser  TGC Cys 560 CAT His	CCA Pro CTC Leu 545 GAG Glu ATT	TGC Cys 530 ATC Ile CGC Arg	CAG Gln GCC Ala TGC Cys	TGT Cys CAC His GGC Gly CAC His 580	GTG Val GTG Val AAG Lys 565 GAC Asp	1806 1854
<i>35</i>	CAG Gln ATG Met CGC Arg 550 AGA Arg	CGG Arg TGC Cys 535 CAG Gln TTC Phe	CAC His 520 GGT Gly CAC His GTC Val	GTC Val AAG Lys ACC Thr CAG Gln	CGC Arg GCC Ala GGG Gly TCC Ser	ATT Ile TTC Phe GAG Glu 5555 AGC Ser	CAC His ACC Thr 540 AAG Lys CAG Gln	ACA Thr 525 CAG Gln CCC Pro TTG Leu	GGT Gly GCC Ala TAC Tyr GCC Ala GTG	GAG Glu AGC Ser GTC Val AAT ASN 575	Ala  AAG Lys  TCC Ser  TGC Cys 560 CAT His	CCA Pro CTC Leu 545 GAG Glu ATT Ile	TGC Cys 530 ATC Ile CGC Arg CGC Arg	CAG Gln GCC Ala TGC Cys CAC His	TGT Cys CAC His GGC Gly CAC His 580	GTG Val GTG Val AAG Lys 565 GAC Asp	1806 1854 1902
<i>35</i>	CAG Gln ATG Met CGC Arg 550 AGA Arg	CGG Arg TGC Cys 535 CAG Gln TTC Phe	CAC His 520 GGT Gly CAC His GTC Val	GTC Val AAG Lys ACC Thr CAG Gln	CGC Arg GCC Ala GGG Gly TCC Ser 570	ATT Ile TTC Phe GAG Glu 5555 AGC Ser	CAC His ACC Thr 540 AAG Lys CAG Gln	ACA Thr 525 CAG Gln CCC Pro TTG Leu	GGT Gly GCC Ala TAC Tyr GCC Ala GTG	GAG Glu AGC Ser GTC Val AAT ASN 575	Ala  AAG Lys  TCC Ser  TGC Cys 560 CAT His	CCA Pro CTC Leu 545 GAG Glu ATT Ile	TGC Cys 530 ATC Ile CGC Arg CGC Arg	CAG Gln GCC Ala TGC Cys CAC His	TGT Cys CAC His GGC Gly CAC His 580	GTG Val GTG Val AAG Lys 565 GAC Asp	1806 1854 1902
<i>35</i>	CAG Gln ATG Met CGC Arg 550 AGA Arg	CGG Arg TGC Cys 535 CAG Gln TTC Phe	CAC His 520 GGT Gly CAC His GTC Val	GTC Val AAG Lys ACC Thr CAG Gln CCA Pro 585	CGC Arg GCC Ala GGG Gly TCC Ser 570 CAC His	ATT Ile TTC Phe GAG Glu 5555 AGC Ser AAG Lys	CAC His ACC Thr 540 AAG Lys CAG Gln TGC Cys	ACA Thr 525 CAG Gln CCC Pro TTG Leu AGC Ser	GCC Ala TAC Tyr GCC Ala GTG Val 590 ATC	GAG Glu AGC Ser GTC Val AAT ASN 575 TGC Cys	Ala  AAG Lys  TCC Ser  TGC Cys 560 CAT His  AGC Ser	CCA Pro CTC Leu 545 GAG Glu ATT Ile AAG Lys	TGC Cys 530 ATC Ile CGC Arg GCC Ala	CAG Gln GCC Ala TGC Cys CAC His TTC Phe 595	TGT Cys CAC His GGC Gly CAC His 580 GTG Val	GTG Val GTG Val AAG Lys 565 GAC ASP	1806 1854 1902
<i>35</i>	CAG Gln ATG Met CGC Arg 550 AGA Arg	CGG Arg TGC Cys 535 CAG Gln TTC Phe	CAC His 520 GGT Gly CAC His GTC Val	GTC Val AAG Lys ACC Thr CAG Gln CCA Pro 585	CGC Arg GCC Ala GGG Gly TCC Ser 570 CAC	ATT Ile TTC Phe GAG Glu 5555 AGC Ser AAG Lys	CAC His ACC Thr 540 AAG Lys CAG Gln TGC Cys	ACA Thr 525 CAG Gln CCC Pro TTG Leu AGC Ser	GCC Ala TAC Tyr GCC Ala GTG Val 590 ATC	GAG Glu AGC Ser GTC Val AAT ASN 575 TGC Cys	Ala  AAG Lys  TCC Ser  TGC Cys 560 CAT His  AGC Ser	CCA Pro CTC Leu 545 GAG Glu ATT Ile AAG Lys	TGC Cys 530 ATC Ile CGC Arg GCC Ala	CAG Gln GCC Ala TGC Cys CAC His TTC Phe 595	TGT Cys CAC His GGC Gly CAC His 580 GTG Val	GTG Val GTG Val AAG Lys 565 GAC ASP	1806 1854 1902 1950

_	TAC Tyr	CTG Leu 615	TGT Cys	GAT Asp	AAG Lys	TGT Cys	GGG Gly 620	CGT Arg	GGC Gly	TTC Phe	AAC Asn	CGG Arg 625	GTA Val	GAC Asp	AAC Asn	CTG Leu	2046
<b>.</b>	Arg 630	Ser	CAC His	Val	Lys	Thr 635	Val	His	Gln	Gly	Lys 640	Ala	Gly	Ile	Lys	11e 645	2094
10	CTG Leu	GAG Glu	CCC Pro	GAG Glu	GAG Glu 650	GGC Gly	AGT Ser	GAG Glu	GTC Val	AGC Ser 655	GTG Val	GTC Val	ACT Thr	GTG Val	GAT Asp 660	GAC Asp	2142
15	ATG Met	GTC Val	ACG Thr	CTG Leu 665	GCT Ala	ACC Thr	GAG Glu	GCA Ala	CTG Leu 670	Ala	GCG Ala	ACA Thr	GCC Ala	GTC Val 675	ACT Thr	CAG Gln	2190
20	CTC Leu	ACA Thr	GTG Val 680	GTG Val	CCG Pro	GTG Val	GGA Gly	GCT Ala 685	GCA Ala	GTG Val	ACA Thr	GCC Ala	GAT Asp 690	GAG Glu	ACG Thr	GAA Glu	2238
	GTC Val	CTG Leu 695	AAG Lys	GCC Ala	GAG Glu	ATC Ile	AGC Ser 700	AAA Lys	GCT Ala	GTG Val	AAG Lys	CAA Gln 705	GTG Val	CAG Gln	GAA Glu	GAA Glu	2286
25	GAC Asp 710	CCC Pro	AAC Asn	ACT Thr	CAC His	ATC 11e 715	CTC Leu	TAC Tyr	GCC Ala	TGT Cys	GAC Asp 720	TCC Ser	TGT Cys	GGG Gly	GAC Asp	AAG Lys 725	2334
30	TTT Phe	CTG Leu	GAT Asp	GCC Ala	AAC Asn 730	AGC Ser	CTG Leu	GCT Ala	CAG Gln	CAT His 735	GTG Val	CGA Arg	ATC Ile	CAC His	ACA Thr 740	Ala	_2382
<i>35</i>	CAG Gln	GCA Ala	CTG Leu	GTC Val 745	ATG Met	TTC Phe	CAG Gln	ACA Thr	GAC Asp 750	GCG Ala	GAC Asp	TTC Phe	TAT Tyr	CAG Gln 755	Gln	TAT Tyr	2430
	GGG Gly	CCA Pro	GGT Gly 760	Gly	ACG Thr	TGG Trp	CCT Pro	GCC Ala 765	GGG Gly	CAG Gln	GTG Val	CTG Leu	CAG Gln 770	GCT Ala	GGG	GAG Glu	2478
40	CTG Leu	GTC Val 775	Phe	CGC	CCT Pro	CGC Arg	GAC Asp 780	Gly	GCT Ala	GAG Glu	GGC Gly	CAG Gln 785	CCC Pro	GCA Ala	CTG Leu	GCA Ala	2526
45	GAG Glu 790	Thr	TCC Ser	CCT Pro	ACA Thr	CCT Pro 795	CCT Pro	GAA Glu	TGT Cys	CCC	CCG Pro 800	CCT Pro	GCC Ala	GAG Glu	TGA	GCTGGCG	2578
	GCC	CTTC	TGA	CTGT	TATT	TT A	AGGA	TGGA	T GG	CACC	CTGG	AAC	CGGG	AAG	GGTG	GCCTGT	2638
50			AGA							AAAA	AAAA	AA					2680
		TATE	ORMA	TAN	711	SEO :	TD N	0 2	•								

5		`	(E	A) LÃ B) AF D) TO	RT: A	mino	säur	e	uren	1						
				r DES												
		(xi)	SEC	QUENZ	BESC	HRE	BUNC	s: SE	EQ II	NO:	2:					
10	Met 1	Asp	Phe	Pro	Gln 5	His	Ser	Gln	His	Val 10	Leu	Glu	Gln	Leu	Asn 15	Gln
15	Gln	Arg	Gln	Leu 20	Gly	Leu	Leu	Суѕ	Asp 25	Cys	Thr	Phe	Val	Val 30	Asp	Gly
	Val	His	Phe 35	Lys	Ala	His	Lys	Ala 40	Val	Ļел	Ala	Ala	Cys 45	Ser	Glu	Tyr
20	Phe	Lys 50	Met	Leu	Phe	Val	Asp 55	Gln	Lys	Asp	Val	Val. 60	His	Leu	Asp	Ile
	Ser 65	Asn	Ala	Ala	Gly	Leu 70	Gly	Gln	Met	Leu	G1u 75	Phe	Met	Tyr	Thr	Ala 80
25	Lys	Leu	Ser	Leu	Ser 85	Pro	Glü	Asn	Val	Asp 90	Asp	Val	Leu	Ala	Val 95	Ala
	Thr	Phe		Gln -100-		Gln		Ile			Ala	Cys	His	Ala 110	Leu	Lys
30	Ser	Leu	Ala 115	Glu	Pro	Ala	Thr	Ser 120	Pro	Gly	Gly	Asn	A1a 125	Glu	Ala	Leu
<b>35</b> .	Ala	Thr 130		Gly	Gly	Asp	Lys 135		Ala	Lys	Glu	Glu 140	Lys	Val	Ala	Thr
	Ser 145	Thr	Leu	Ser	Arg	Leu 150	Glu	Gln	Ala	Gly	Arg 155	Ser	Thr	Pro	Ile	Gly 160
40		Ser			165					170					175	
	Ala	Ser	Gly	Ala 180		Gln	Thr	Glu	Lys 185		Asp	Ala	Pro	Arg 190	Glu	Pro
45	Pro	Pro	Val 195	Glu	Leu	Lys	Pro	Asp 200	Pro	Thr	Ser	Gly	Met 205	Ala	Ala	Ala
	Glu	Ala 210	Glu	Ala	Ala	Leu	Ser 215	Glu	Ser	Ser	Glu	G1n 220		Met	Glu	Val
50	Glu 225	Pro	Ala	Arg	Lys	Gly 230	Glu	Glu	Glu	Gln	Lys 235	Glu	Gln	Glu	Glu	Gln 240

	Glu	Glu	Glu	Gly	Ala 245	Gly	Pro	Ala	Glu	Val 250	Lys	Glu	Glu	Gly	Ser 255	Gln
5	Leu	Glu	Asn	Gly 260	Glu	Ala	Pro	Glu	G1u 265	Asn	Glu	Asn	Glu		Ser	Ala
	Gly	Thr	Asp 275	Ser	Gly	Gln	Glu	Leu 280	Gly	Ser	Glu	Ala	Arg 285	Gly	Leu	Arg
10	Ser	Gly 290	Thr	Tyr	Gly	Asp	Arg 295	Thr	Glu	Ser	Lys	A1a 300	Tyr	Gly	Ser	Val
15	Ile 305	His	Lys	Суѕ	Glu	Asp 310	Cys	Gly	Lys	Glu	Phe 315	Thr	His	Thr	Gly	Asn 320
		Lys			325				•	330					335	
20		Glu		340					345					350		
			355					360					365			Gly
25	Lys	Ser 370	Tyr	Arg	Leu	Ile	Ser 375	Fén	Leu	Asn	Leu	His 380	ГÀЗ	ГÀЗ	Arg	His
		-Gly	-Glu	_Ala	Arg		Arg	Суз	Glu	Asp	Cys	Gly	Lys	Leu	Phe	Thr
30	385					390					395					400
30	385 Thr	Ser	Gly	Asn	Leu 405	390	Arg	His	G1n	Leu 410	395 Val	His	Ser	Gly	Glu 415	Lys
30	385 Thr Pro	Ser	Gly Gln	Asn Cys 420	Leu 405 Asp	390 Lys Tyr	Arg Cys	His	Gln Arg 425	Leu 410 Ser	Val	His Ser	Ser	Gly Pro 430	Glu 415 Thr	Lys Ser
	385 Thr Pro	Ser Tyr Met	Gly Gln Arg 435	Asn Cys 420 His	Leu 405 Asp	J90 Lys Tyr Glu	Arg Cys Thr	His Gly His	Gln Arg 425 Asp	Leu 410 Ser Thr	395 Val Phe Asp	His Ser Lys	Ser Asp Glu 445	Gly Pro 430 His	Glu 415 Thr	Lys Ser Cys
	385 Thr Pro Lys	Ser Tyr Met His	Gly Gln Arg 435 Cys	Asn Cys 420 His	Leu 405 Asp Leu	190 Lys Tyr Glu Lys	Arg Cys Thr Phe 455	His Gly His 440 Asn	Gln Arg 425 Asp Gln	Leu 410 Ser Thr	795 Val Phe Asp	His Ser Lys Asn 460	Ser Asp Glu 445 Leu	Gly Pro 430 His	Glu 415 Thr Lys	Lys Ser Cys
35 40	Thr Pro Lys Pro Leu 465	Tyr Met His 450	Gly Gln Arg 435 Cys	Asn Cys 420 His Asp	Leu 405 Asp Leu Lys	390 Lys Tyr Glu Lys Ala 470	Arg Cys Thr Phe 455 Asp	His Gly His 440 Asn	Gln Arg 425 Asp Gln Pro	Leu 410 Ser Thr Val	Phe Asp Gly Lys 475	His Ser Lys Asn 460	Asp Glu 445 Leu Arg	Gly Pro 430 His Lys	Glu 415 Thr Lys Ala	Lys Ser Cys His Gly 480
35	Thr Pro Lys Pro Leu 465	Tyr Met His 450 Lys	Gly Gln Arg 435 Cys Ile	Asn Cys 420 His Asp	Leu 405 Asp Leu Lys Ile Thr 485	390 Lys Tyr Glu Lys Ala 470 Ser	Arg Cys Thr Phe 455 Asp	His Gly His 440 Asn	Gln Arg 425 Asp Gln Pro	Leu 410 Ser Thr Val Leu Lys 490	Phe Asp Gly Lys 475	His Ser Lys Asn 460 Cys	Asp Glu 445 Leu Arg	Pro 430 His Lys	Glu 415 Thr Lys Ala Cys	Lys Ser Cys His Gly 480
35 40	Thr Pro Lys Pro Leu 465 Lys	Ser Tyr Met His 450 Lys Gln	Gly Gln Arg 435 Cys Ile Phe	Asn Cys 420 His Asp His Thr	Leu 405 Asp Leu Lys Ile Thr 485 Pro	390 Lys Tyr Glu Lys Ala 470 Ser	Arg Cys Thr Phe 455 Asp Gly Val	His Gly His 440 Asn Gly Asn	Gln Arg 425 Asp Gln Pro	Leu 410 Ser Thr Val Leu Lys 490 His	Phe Asp Gly Lys 475	His Ser Lys Asn 460 Cys	Asp Glu 445 Leu Arg	Gly Pro 430 His Lys Glu Arg	Glu 415 Thr Lys Ala Cys Ile 495	Lys Ser Cys His Gly 480 His

		Pro	Cys 530	Gln	Cys	Val	Met	Cys 535	Gly	Lys	Ala	Phe	Thr 540	Gln	Ala	Ser	Ser
į		Leu 545	Ile	Ala	His	Val	Arg 550	Gln	His	Thr	Gly	Glu 555	Lys	Pro	Tyr	Val	Cys 560
		Glu	Arg	Cys	Gly	Lys 565	Arg	Phe	Val	Gln	Ser 570	Ser	Gln	Leu	Ala	Asn 575	His
	10 .	Ile	Arg	His	His 580	Asp	Asn	Ile	Arg	Pro 585	His	Lys	Cys	Ser	Val 590	Cys	Ser
	15	Lys	Ala	Phe 595	Val	Asn	Val	Gly	Asp 600	Leu	Ser	Lys	His	I1e 605	Ile	Ile	His
		Thr	Gly 610	Glu	Lys	Pro	Tyr	Leu 615	Cys	Asp	Lys	Cys	G1y 620	Arg	Gly	Phe	Asn
	20	Arg 625	Val	Asp	Asn	Leu	Arg 630	Ser	His	Val	Lýs	Thr 635	Val	His	Gln	Gly	Lys 640
		Ala	Gly	Ile	Lys	Ile 645	Leu	Glu	Pro	Glu	G1u 650	Gly	Ser	Glu	Val	Ser 655	Val
_	25	Val	Thr	Val	Asp 660	Asp	Met	Val	Thr	Leu 665	Ala	Thr	Glu	Ala	Leu 670	Ala	Ala
	30	Thr	Ala	Val 675	Thr	Gln	Leu	Thr	Val 680	Val	Pro	Val	-G <del>ly</del> -	A1a 685	Ala	Val	Thr
			690					695					700				Lys
	35	705			•		710					715					Asp 720
-						725					730					7.35	Val
	40				740					745					750		Asp
	<b>4</b> 5			755					760					765			Val
		•	770					775			٠		780				Gly
	50	Gln 785	Pro	Ala	Leu	Ala	Glu 790	Thr	Ser	Pro	Thr	Pro 795	Pro	Glu	Cys	Pro	800

### Patentansprüche

10

30

35

40

- Isoliertes Protein mit der in SEQ ID NO:2 dargestellten Aminosäuresequenz sowie die daraus durch Substitution, Insertion oder Deletion von einem oder mehreren Aminosäureresten erhältlichen Muteine, die noch die wesentlichen biologischen Eigenschaften des in SEQ ID NO:2 dargestellt in Proteins besitzen.
- 2. Protein gemäß Anspruch 1, dadurch gekennzeichnet, daß es sich um ein humanes Protein handelt.
- 3. Nukleinsäuresequenz codierend für ein Protein gemäß Anspruch 1.
- Nukleinsäuresequenz nach Anspruch 3, dadurch gekennzeichnet, daß sie für ein Protein codiert, das mindestens 95 % Identität mit der in SEQ ID NO:2 dargestellten Sequenz besitzt.
- Nukleinsäuresequenz nach Anspruch 3, dadurch gekennzeichnet, daß sie die in SEQ ID NO:1 dargestellte Struktur
   besitzt.
  - 6. Vektor enthaltend eine Nukleinsäuresequenz gemäß Anspruch 3 5, funktionell verknüpft mit mindestens einem Regulationselement.
- Wirtsorganismus, transformiert mit einer Nukleinsäuresequenz nach Anspruch 3.
  - 8. Wirtsorganismus, transformiert mit einem Vektor gemäß Anspruch 6.
- Verfahren zur Herstellung eines Proteins gemäß Anspruch 1, dadurch gekennzeichnet, daß man einen Wirtsorganismus gemäß Anspruch 6 unter Bedingungen kultiviert, die die Expression des Proteins erlauben und anschließend das exprimierte Protein vom Wirtsorganismus abtrennt und in reiner Form isoliert.
  - 10. Verwendung eines Proteins gemäß Anspruch 1 zur Identifizierung von spezifischen transkriptionsmodulierenden Substanzen.
  - 11. Verfahren zur Identifizierung von spezifischen transkriptionsmodulierenden Substanzen, das folgende Schritte
    - (a) Inkubation des Proteins gemäß Anspruch 1 mit dem Genprodukt von myc unter Bedingungen, unter denen sich ein Proteinkomplex zwischen diesen beiden Proteinen ausbildet,
    - (b) Inkubation der beiden Proteine unter ansonst gleichen Bedingungen wie (a) jedoch in Anwesenheit einer oder mehrerer Substanzen, die auf spezifische transkriptionsmodulierende Aktivitäten zu testen sind,
    - (c) Ermitteln des Unterschiedes in der Proteinkomplexbildung zwischen (b) und (a),
      - (d) Auswahl solcher Substanzen, bei denen gemäß Schritt (b) eine andere Proteinkomplexbildung erhalten wurde als bei Schritt (a).
- 45 12. Verwendung eines Proteins gemäß Anspruch 1 als Antigen zur Herstellung von spezifischen Antikörpern.
  - 13. Verwendung einer Nukleinsäuresequenz nach Anspruch 3 zur Gentherapie.
  - 14. Verwendung einer zu der Sequenz gemäß Anspruch 3 komplementären Nukleinsäuresequenz zur Gentherapie.
- Verwendung nach Anspruch 13 oder 14, dadurch gekennzeichnet, daß man durch die exogen zugeführte Nukleinsäuresequenz die zelluläre Konzentration des Proteins gemäß Anspruch 1 erhöht oder erniedrigt.



**Europäisches Patentamt** 

**European Patent Office** 

Office uropéen des brevets



11) EP 0 875 567 A3

(12)

## **EUROPÄISCHE PATENTANMELDUNG**

(88) Veröffentlichungstag A3: 22.12.1999 Patentblatt 1999/51

(43) Veröffentlichungstag A2: 04.11.1998 Patentblatt 1998/45

(21) Anmeldenummer: 98106426.4

(22) Anmeldetag: 08.04.1998

(51) Int. Cl.<sup>6</sup>: **C12N 15/12**, C07K 14/47, C12N 15/63, C12N 1/21, G01N 33/68, C07K 16/18, A61K 48/00

(84) Benannte Vertragsstaaten:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU
MC NL PT SE
Benannte Erstreckungsstaaten:
AL LT LV MK RO SI

(30) Priorität: 30.04.1997 DE 19718249

(71) Anmelder: BASF AKTIENGESELLSCHAFT 67056 Ludwigshafen (DE) (72) Erfinder:

- Peukert, Karen
   35094 Lahntal-Sterzhausen (DE)
- Haenel, Frank, Dr. 07745 Jena (DE)
- Eilers, Martin, Prof. Dr. 35043 Marburg-Cappel (DE)

(54) Myc-bindende Zinkfinger-Proteine, ihre Herstellung und ihre Verwendung

(57) Myc-bindende Zinkfingerproteine, ihre Herstellung und ihre Verwendung.



# Europäisches EUROPÄISCHER TEILRECHERCHENBERICHT

Nummer der Anmeldung

der nach Regel 45 des Europäischen Patentübereinkommens für das weitere Verfahren als europäischer Recherchenbericht gilt

EP 98 10 6426

	EINSCHLÄGIGE	DOKUMENTE		· · · · · · · · · · · · · · · · · · ·
Kategorie	Kennzeichnung des Dokum der maßgebliche	ents mit Angabe, soweit erforderlich en Teile	Betrifft Anspruch	KLASSIFIKATION DER ANMELDUNG (Int.C1.6)
D,X	THOMAS C. SCHULZ ET arrangement of 13 z vertebrate gene Z13' BIOCHEMICAL JOURNAL Bd. 311, Nr. 1, 1. Oktober 1995 (199219-224, XP002117529 das ganze Dokument	inc fingers in the , , 95-10-01), Seiten	1-3, 6-10, 12-15	C12N15/12 C07K14/47 C12N15/63 C12N1/21 G01N33/68 C07K16/18 A61K48/00
X	mapping of 16 novel finger-encoding cDN candidate genes for malignant disorders GENOMICS, Bd. 27, Nr. 2, 20. Seiten 259-264, XPO * Zusammenfassung;	As identify putative developmental and " Mai 1995 (1995-05-20) 02117526 Tabelle 1 *	6-10, 12-15	
	-& R59u022 Datenban	k Hs20647		
	Zugriffsnummer U206 TOMMRUP N. ET AL.:" protein (ZNF151) mR	Human zinc finger		RECHERCHIERTE SACHGEBIETE (Int.Cl.5)
	XP002117530			C12N
		-/		CO7K   GO1N   A61K
UNVC	LLSTÄNDIGE RECHE	RCHE		
in alnom	nerchenabtellung ist der Auffassung, de solichen Umfang nicht emspricht bzw. nik für diese Ansprüche nicht, bzw. nur	iß ein oder mehrere Ansprüche, den Vor entsprechen, daß sinnvolle Ermittlungen r teilweise, möglich sind.	rschriften des EPÜ uber den Stand	
	lig recherchierte Patentansprüche:	. *		
Unvollstä	indig recherchierte Patentansprüche:			
Nicht rec	herchierte Patentansprüche:			
	r die Beschränkung der Recherche:			
Beh bez dur	andlung des menschli 1ehen (Artikel 52(4)	-15 sich auf ein Verf chen/tierischen Körpe EPÜ), wurde die Rech te sich auf die angef g/Zusammensetzung.	ers nerche	
	Recherchenort	Abschlußdatum der Recherche	· -	Pr@fer
	DEN HAAG	5. Oktober 199	9 Moi	ntero Lopez, B
X:vo: Y:vo an A:tec O:nic	KATEGORIE DER GENANNTEN DOK n besonderer Bedeutung allein betrach n besonderer Bedeutung in Verbindum deren Veröffentlichung derseiben Kate- chnologischer Hintergrund chtschriftliche Offenbarung vischentliteratur	tet E : ålteres Pate nach dem A g mit einer D : in der Anne gorie L : aus anderei	entdokument, das jed unmeldedatum veröff eldung angeführtes D n Gründen angeführt	entlicht worden ist ookument



## EUROPÄISCHER TEILRECHERCHENBERICHT

Nummer der Anmeldun

EP 98 10 6426

	EINSCHLÄGIGE DOKUMENTE		KLASSIFIKATION DER ANMELDUNG (Int.CI.6)
ategorie	Kennzeichnung des Dokuments mit Angabe, soweit erforderlich der maßgeblichen Teile	Betrifft Anspruch	
Α, Ο	ANGELIKA PHILIPP ET AL.: "Repression of Cyclin D1: a novel function of MYC" MOLECULAR AND CELLULAR BIOLOGY, Bd. 14, Nr. 6, Juni 1994 (1994-06), Seiten 4032-4043, XP002117527 * Zusammenfassung * * Seite 4039, linke Spalte, Absatz 2 - Seite 4041, rechte Spalte, letzter Absatz *	1-15	
>,χ	PEUKERT K ET AL: "An alternative pathway for gene regulation by Myc." EMBO JOURNAL, (1997 SEP 15) 16 (18) 5672-86., XP002117528 * das ganze Dokument *	1-15	RECHERCHIERTE SACHGEBIETE (Int.Cl.6)
Ρ,Χ	SCHNEIDER A ET AL: "Association of Myc with the zinc-finger protein Miz -1 defines a novel pathway for gene	1-15	·
	regulation by Myc." CURRENT TOPICS IN MICROBIOLOGY AND IMMUNOLOGY, (1997) 224 137-46., XP002117529 * das ganze Dokument *		